SOCCA: A Security-Oriented Cyber-Physical Contingency Analysis in Power Infrastructures

Samon Zonouz Member, IEEE, Charles M. Davis Member, IEEE, Katherine R. Davis Member, IEEE, Robin Berthier, Rakesh B. Bobba Member, IEEE, William H. Sanders Fellow, IEEE

Abstract—Contingency analysis is a critical activity in the context of the power infrastructure because it provides a guide for resiliency and enables the grid to continue operating even in the case of failure. In this paper, we augment this concept by introducing SOCCA, a cyber-physical security evaluation technique to plan not only for accidental contingencies but also for malicious compromises. SOCCA presents a new unified formalism to model the cyber-physical system including interconnections among cyber and physical components.

The cyber-physical contingency ranking technique employed by SOCCA assesses the potential impacts of events. Contingencies are ranked according to their impact as well as attack complexity. The results are valuable in both cyber and physical domains. From a physical perspective, SOCCA scores power system contingencies based on cyber network configuration, whereas from a cyber perspective, control network vulnerabilities are ranked according to the underlying power system topology.

Index Terms—Cyber-physical systems, security, contingency analysis, situational awareness, state estimation.

I. INTRODUCTION

State estimation and contingency analysis are two of the most fundamental tools for monitoring the power system. State estimation is the process of fitting data from sensors in the field to a system model and determining an estimate of the power system state [1]. State estimation engines use techniques such as the weighted least squares (WLS) algorithm to determine the state of the system based on the measurements [2]. Once the state estimator program determines a system estimate, the estimate is used to run a series of “what if” scenarios referred to as contingency analysis. Contingency analysis performs a series of power flow studies with various pieces of equipment outaged in the model, allowing operators to predict the state of the system for such an event [3].

By its nature, state estimation depends on the communication infrastructure, commonly called the SCADA (supervisory control and data acquisition) system. These systems are currently undergoing many upgrades as part of the smart grid initiative. This initiative does not only affect the SCADA system but also brings the telecommunication revolution to the entire energy delivery infrastructure, from control centers to generation, transmission and distribution substations, and even to customer homes. A direct consequence has been a significant increase in the number of inter-connected cyber components. Manufacturers commonly add ethernet or radio communication modules to controllers, relays, and sensors, along with information report and configuration functionalities such as embedded web servers.

The increased cyber connectivity of the infrastructure and the interdependency of cyber and physical components introduces a greater level of complexity, and securing power system operations against malicious compromise becomes more challenging. Indeed, contingency analysis in the highly interconnected grid should be expanded to include incidents of intentional nature such as cyber attacks.

While the problem of detecting and mitigating cyber intrusions has been extensively studied over the past two decades in the context of traditional IT systems, the requirements and constraints of the smart grid environment in terms of security are different and usually more stringent. For example, power grid components commonly have timing requirements that prevent traditional security solutions from being deployed. The dependencies due to cyber-physical interactions in the grid are not yet well understood. Recently, there have been several attempts to model and analyze the cyber-physical threats in an offline manner [4]–[7]. Zonouz et al. [8] proposed an online framework that fuses uncertain information from distributed power system meters and cyber-side intrusion detectors to detect malicious activities within the cyber-physical system. However, to the best of our knowledge, there has been no efficient online solution proposed for contingency analysis that considers cyber adversary induced physical contingencies.

This paper introduces a cyber-physical contingency analysis framework for analyzing the physical impacts resulting from compromise in the cyber network. In particular, we present Security-Oriented Cyber-Physical Contingency Analysis framework or SOCCA, which takes into account cyber- and power-side network topologies, malicious cyber asset compromises, and power component outages. During an offline process, SOCCA analyzes the cyber network topology and the network firewall rules to automatically generate a network connectivity map as a directed graph encoding accessibility among different hosts (e.g., computers and other cyber components). Then, SOCCA uses the connectivity map to create a partial Markovian state-based model of the power-grid in an online manner. At any time instance, SOCCA estimates the current security state using the generated model and the set of cyber intrusion detection sensor alerts. Using a new cyber-physical security index and cyber-physical state notion, SOCCA measures the criticality level of each system state and produces a ranked list of potential cyber and/or physical contingencies that need to be addressed.

Saman Zonouz s.zonouz@miami.edu is an Assistant Professor in the Electrical and Computer Engineering Department at the University of Miami. Matt Davis and Kate Davis (matt, kate)@powerworld.com are with PowerWorld Corporation. Robin Berthier, Rakesh Bobba, and William H. Sanders (rbobba, whs)@illinois.edu, are Research Scientist, Assistant Research Professor, and Donald Biggar Willett Professor, respectively, in the Electrical and Computer Engineering Department at the University of Illinois.
The proposed framework does not require any additional measurements than those already present in modern control centers and does not impose additional communication requirements. Specifically, cyber intrusion detection systems are already in use in modern control networks, and the sensor alerts from those systems are typically available from a network Security Event Manager (SEM) used to aggregate and manage security alerts. Similarly, the needed power system state information is available from the Energy Management System (EMS).

The contributions of this paper are threefold:

- **Cyber-physical system formalism and automated model generation**: We propose a new formulation of cyber-physical failures and compromises and an algorithm to automatically generate the corresponding models.
- **Cyber-physical security index for the power grid infrastructures**: We present a novel and scalable security index for power grid contingency screening.
- **Cyber-aware contingency analysis**: We present a framework for power grid contingency screening that uses the security index to determine impact and criticality for each state and then ranks contingencies that could be caused by cyber attacks.

SOCCA is not designed to replace the traditional power system contingency analysis solutions, which analyze accidental failures that could occur at any part of the power system due to natural causes. Instead, SOCCA presents a complementary framework that concentrates on potential contingencies due to remote malicious attacks.

This paper is organized as follows. Section II reviews the high-level architecture of the SOCCA framework and how its components are logically interconnected. Section III discusses how SOCCA models various security incidents and their potential correlations. Section IV presents how the cyber-physical model probabilistically determines the current state given the real-time sensory information. Section V describes the way SOCCA explores and analyzes potential cyber-physical contingencies according to the current cyber-physical state estimate and provides a risk-based ranked list.

### II. BACKGROUND

Power system modeling is used extensively in grid operations. The most common types of models used are steady state power flow models and sensitivities. These models are used to monitor the state of the system and predict the effects of changes.

Steady state power system modeling consists of enforcing the conservation of power. Given a set of power injections and withdrawals, the power flow finds the set of voltages and angles that satisfy power balance. The system state may be written as

\[
x = [V, \theta]
\]

where \(V\) is a vector of voltages, \(\theta\) is a vector of voltage angles. The vector of real power loads is \(P_l\) and the vector of reactive power loads is \(Q_l\). Since generator outputs are controllable (within limits), they are collected separately in a vector of controls, \(u\).

The power flow problem can now be written as

\[
f(x, u) = 0
\]

where \(f(x, u) = 0\) is a complex vector representing the injection at each node in the system. The function \(f(x, u)\) represents the system model. It encapsulates factors like line impedances and system topology. Breaking \(f(x, u)\) into real and reactive parts gives

\[
f_i^p = -P_i^p + P_i^l + \sum_{k \in C} |V_i||V_k| (G_{ik} \cos \theta_{ik} + B_{ik} \sin \theta_{ik})
\]

\[
f_i^q = -Q_i^p + Q_i^l + \sum_{k \in C} |V_i||V_k| (G_{ik} \sin \theta_{ik} - B_{ik} \cos \theta_{ik})
\]

These equations represent the nonlinear problem that is commonly called the power flow in power systems literature. The power flow is at the heart of most power systems analysis. It provides the basis for many tools and sensitives that are used to predict the state of the system in the event of an outage.

Because the power flow is a non-linear problem, it is typically solved using the Newton-Raphson method, an iterative technique that requires multiple evaluations and factorizations of a large sparse matrix of sensitivities, the Jacobian matrix \(J\). The repeated factorizations can be a time consuming process, so more efficient approximate methods have been developed. These methods involve applying assumptions to the power flow equations to arrive at a simplified system model [9].

A commonly used simplification reduces the power flow to a linear problem, commonly called the DC power flow in the power system literature [3]. A constant matrix relates system angles and power injections. The DC power flow is the basis for many sensitivities. For example, power transfer distribution factors (PTDFs) estimate the changes in flow due to a transfer across the power system, and line outage distribution factors (LODFs) estimate the changes in flow on a line caused by the outage of another line [3]. PTDFs and LODFs are frequently used to predict the state of the system after an outage [10], [11]. There are also efficient extensions of LODFs to calculate changes in flow due to multiple outages [12].

### III. CYBER-PHYSICAL SYSTEM SECURITY FORMALISM

In this section, we explain how we model power grid security attacks using a stochastic Markovian mechanism, and how models are automatically generated given the power grid cyber-physical topology.

#### A. Cyber-Physical State Notion

Before discussing how the current state of the power grid will be represented in SOCCA, a concise cyber-physical state notion needs to be defined. The security state, \(s\), is the set of privileges that an adversary (or group of adversaries) has obtained out of the domain of possible privileges, \(D\), which encompasses all privileges on all cyber hosts in the system as well as all connected physical power system devices.

The security state is used to indicate whether or not a specific event (contingency\(^1\)) has occurred in the grid infrastructure. In particular, we consider two types of contingencies.

\(^1\)In this paper, we treat both cyber- and power side incidents as contingencies.
First, there are cyber-side vulnerability exploitations, which are carried out by an attacker to obtain specific privileges and improve his or her control over the power network. Therefore, the information in a state denotes the attacker’s privileges in that state, e.g., root access on a mission-critical host system in power control room. Those privileges are used to determine what further malicious damage the attacker can cause in that state. Second, there are consequences, which are caused by the adversary after he or she obtained the required privileges. Specifically, consequences are defined to be incidents which affect the physical operation of the underlying power system. As a case in point, a transmission line outage, whether due to lightning or a remote malicious “open” command to a power relay, is a power-side consequence which results in a redistribution of power flow. The intent is to always operate the system such that the redistribution of power flow does not affect the end-user consumers.

B. Modeling Power Failures and Cyber Compromises

Generally, every power grid attack path consists of an escalating series of malicious actions by the adversary. The system’s initial state is (Ø), in which no contingency has occurred and the attacker does not yet have privilege in the system. Starting from this state, the adversary aims to gain the set of privileges required to reach his or her goals, e.g., causing a power transmission line outage by opening the corresponding relay.

More specifically, every cyber-physical attack is in a finite set of security states \( S \) that cover all possible security conditions that the system could be in. The system is in one of the security states \( s \) at each time instant. From the system’s current state \( s \), there are two types of transitions, corresponding to 1) adversarial vulnerability exploitations, and 2) malicious power contingencies. Formally, the attacker can choose and take an adversarial action \( a \in A \) admissible in \( s \), resulting in a state transition to \( s' \).

To enumerate all possible attack scenarios, we model the adversarial actions as a discrete Markov decision process (MDP) [13]. A discrete Markovian decision process \( \Gamma \) is defined as a tuple \( (S,A,F,\cdot,P,\gamma) \) where \( S \) is the security state space, assumed to be an arbitrary non-empty set endowed with the discrete topology. \( A \) is the set of actions which consists of adversarial vulnerability exploitations. For every \( s \in S \), \( A(s) \subseteq A \) is the set of admissible actions at state \( s \). The measurable function \( F : S \rightarrow \mathbb{R} \) is the susceptibility measure to attacks calculated for each state, and \( P \) is the transition probability function. That is, if the present state of the system is \( s \in S \) and the attacker takes an action \( a \in A(s) \), resulting in state transition to state \( s' \) with probability \( P(s'|s,a) \), he or she obtains an immediate reward of \( F(s') \). The discounting factor is \( \gamma \) which is normalized, i.e., \( 0 < \gamma < 1 \). The discounting factor in control theory is a coefficient that models the fact that a future reward is worth less than the same amount of immediate reward.

C. Automatic MDP Generation

SOCCA automatically generates the MDP model for the power network given the control network topology, access control policies, and cyber-physical interconnections within the power grid. The power network’s access control policies, such as firewall rule sets, are composed of rules about sources (IP/port addresses) that are either allowed or not allowed to reach a destination. SOCCA parses the rulesets and creates a binary network connectivity matrix that is a Cartesian product of host systems. The \([i,j]\) entry of the matrix takes on a true value if traffic from host \( h_i \) to host \( h_j \) is allowed, and a false value otherwise. The connectivity matrix always includes an Internet node representing a group of hosts outside of the network where attackers are assumed to initially reside.

The connectivity matrix incorporates all the possible accesses allowed by the global access policies. However, an attacker, with control over a particular host computer in the network, needs a vulnerability in one of the accessible hosts to exploit and improve his or her privileges. Therefore, SOCCA further refines the matrix to encode only the adversarial paths that can occur through vulnerability exploitations. In particular, SOCCA analyzes the power grid topology input to find the set of known system vulnerabilities in host systems. Given individual vulnerabilities, SOCCA determines their difficulty level through the web-based national vulnerability database [14] that uses the common vulnerability scoring system [15]. SOCCA converts the reported high, medium, low format to numerical 0.25, 0.5, 0.75 values that represent the attacker’s success rates. Consequently, the individual nonzero values in the connectivity matrix are updated accordingly, i.e., SOCCA replaces them either with zero if the corresponding hosts are not vulnerable or with relevant success rates otherwise.

Note that, while we use specific numerical values for transition probabilities in the rest of this work to demonstrate and evaluate SOCCA, a sensitivity analysis can be undertaken by varying the transition probabilities to identify structural weaknesses that are agnostic to transition probabilities. Similarly, while we focus on transition probabilities based on known vulnerabilities, the framework is flexible and zero-day attacks and vulnerabilities can be accounted for by using a small non-zero transition probability on all allowed connections.

To generate the MDP model, SOCCA traverses the connectivity matrix and concurrently updates the model following an incremental process. First, the MDP’s initial state \( (Ø) \) is created and the MDP generation starts with the network’s entry point (Internet) node in the connectivity matrix. Considering
the connectivity matrix as a weighted directed graph, a depth-first search (DFS) is run on the graph. While the search is recursively traversing the graph, it keeps track of the current state in the MDP, i.e., the set of privileges already gained through the path traversed so far. When the search meets a graph edge \([i, j]\) that crosses over privilege domains \(h_i\) to \(h_j\), a state transition \(a_{ij} \in A\) in the MDP is created if the current state in the MDP does not include the privilege domain of the host to which the edge leads, i.e., \(h_j\). The transition in the MDP is between the current state and the state that includes exactly the same privilege set as the current state plus the host \(h_j\) directed by the graph edge \([i, j]\). The transition created is labeled with the appropriate success rate (transition probability) from the connectivity matrix. The MDP’s current state in the algorithm is then updated to the latter state, and the algorithm proceeds until no further updates to the MDP are possible according to the connectivity matrix. Figure 1 shows a simplified power network with vulnerabilities and its corresponding MDP model. Connectivity matrix elements are denoted as red arrows among pairs of network components.

It is noteworthy that SOCCA’s MDP generation engine takes the known system vulnerabilities into consideration while generating the models. To address possible vulnerabilities targeted by zero-day attacks, the engine can incorporate those vulnerabilities to facilitate worst-case system contingency analysis. Additionally, the algorithm can handle generation of MDP models for systems where an adversary or group of adversaries starts from the Internet and can penetrate into the network from different vulnerable entry points. Several attackers would be modeled as a single “more powerful” attacker who can penetrate from any of those entry points.

Finally, SOCCA enhances the MDP model to also consider power contingencies using the power grid topology input that encodes the cyber-power interconnections, i.e., which power components are controlled by a particular host. In particular, considering every MDP state and the attacker’s privileges in that state, SOCCA determines whether any malicious power contingency could occur and creates the required states and transitions accordingly. Those transitions will have the success rates of 1.0, because once the attacker gains the required privileges, he or she can cause a power contingency, by directly sending the corresponding command and without exploiting a vulnerability.

IV. CYBER-PHYSICAL SECURITY INDEX

SOCCA uses a new power grid security index to evaluate the security level of each MDP state. The proposed index takes into consideration the severity of the potential malicious physical consequences (i.e., percentage of line overload) and the difficulty to penetrate into the power network. In particular, using a defense-centric metric, SOCCA measures how susceptible the power system is to cyber attack induced contingencies, e.g., line outages, at every MDP state. As explained in the previous section, states capture the set of contingencies that can occur due to malicious actions. Then, SOCCA makes use of an adversary-driven metric to quantify how the attackers can obtain the required privileges to cause those physical contingencies. Note that the physical contingencies considered here are steady-state contingencies. Although not discussed, SOCCA could potentially be modified to study transient contingencies as well.

Here, to measure the power system’s susceptibility to cyber attack induced contingencies (e.g., line outages) for each state, SOCCA updates the admittance matrix according to the line outages encoded in that state, and solves the AC power flow equations using the iterative Newton-Raphson algorithm to calculate the line flows. Alternatively, an approximate DC model may be used, requiring only a linear solution, but potentially sacrificing detail [16]. SOCCA estimates the susceptibility degree using a modified version of the performance index [12] that assigns 0 to a state if there are no line flow violations and a positive value otherwise, computed using the following equation:

\[
F(s) = \sum_{l \in L} \left[ \max\left(\frac{f_0(l)}{f_{\text{MAX}}(l)} - 1, 0\right) \right]^2. \tag{5}
\]

Here, \(L\) is the set of all lines, \(f_0(l)\) denotes flow on line \(l\) in state \(s\), and \(f_{\text{MAX}}(l)\) denotes the maximum flow allowed on line \(l\). In the event that a power flow fails to converge, a severe physical impact can be assumed which justifies setting \(F(s)\) to a large number, the outage severity. The outage severity should be much larger than any of the line severity measures.

To calculate the overall cyber-physical security index for each state \(I : s \rightarrow \mathbb{R}\), SOCCA uses the above power system performance index in the following dynamic programming equation:

\[
I(s) = \max_{a \in A(s)} \left\{ \gamma \sum_{s' \in S} P(s'|s, a)[\Delta F(s, s') + I(s')] \right\}, \tag{6}
\]

where \(I(.)\) (as the MDP’s value function) is formulated as a function of difficulty levels of various cyber attack paths \(P(s'|s, a)\) and their physical impact \(\Delta F(s, s') = F(s') - F(s)\). To solve Equation (6), SOCCA implements the value iteration algorithm. As formulated, proximity of the attacker to a physical component, existence of easy-to-traverse attack paths, and high final physical impacts of the attacks improve the fitness of a system state from the point of view of the adversary.

V. CONTINGENCY SCREENING

Algorithm 1 shows the pseudo-code for the algorithm that SOCCA employs to evaluate and rank individual cyber-physical contingencies. Briefly, SOCCA calculates an ordered list of single contingencies first, and investigates multiple contingencies ordered by adversarial preferences until either all the contingencies are analyzed or a predefined deadline has passed.

The cyber-physical contingency selection algorithm receives the generated MDP, the current state of the power grid and a hard deadline for the online analysis, and returns the contingency list (Algorithm 1). SOCCA starts off with initializing a temporal buffer list and a FIFO (first in, first out) queue (Lines 1, 2). Each state is assigned an initial color value (white, indicating the state has not yet been checked), as well as the corresponding physical performance and cyber-physical security indices as discussed in Section IV (Lines 3-6). At any
Algorithm 1: Cyber-Physical Contingency Selection

Inputs: MDP, current state, deadline
Output: [Contingency List]

1. List $B \leftarrow \emptyset$;
2. Queue $Q \leftarrow \emptyset$;
3. for $s \in S$ do
   4. $\text{Color}[s] \leftarrow \text{White}$;
   5. $F(s) \leftarrow \sum_{l \in L} \max \left\{ \frac{f_{l}(t)}{\max \{f_{l}(t)\}} - 1, 0 \right\}^2$;
   6. $I(s) \leftarrow \max_{a \in \mathcal{A}(s)} \left\{ \gamma \sum_{s' \in S} P(s'|s,a)[\Delta F(s,s') + I(s')] \right\}$;
4. end
5. $\text{Color}[\text{current state}] \leftarrow \text{Gray}$;
6. $\text{Enqueue}(Q, \text{current state})$;
7. while $(\text{get_current_state}) \leq \text{deadline}$ and $(Q \neq \emptyset)$ do
   8. $s \leftarrow \text{Dequeue}(Q)$;
   9. for $a \in A(s)$ do
      10. $R(s,a) \leftarrow \sum_{s' \in S} P(s'|s,a)[\Delta F(s,s') + I(s')]$;
      11. Insert($B, [R(s,a), s, a]$);
      12. end
   13. $\text{Sort}(B)$;
   14. $\text{Concatenate}(\text{Contingency List}, B)$;
   15. for $b \in B$ do
      16. if $\text{Color}[b_{s,b,a}] = \text{White}$ then
         17. $\text{Color}[b_{s,b,a}] \leftarrow \text{Gray}$;
         18. $\text{Enqueue}(Q, s'_{b,s,b,a})$;
      19. end
   20. $B \leftarrow \emptyset$;
   21. $\text{Color}[s] \leftarrow \text{Black}$;
22. end
23. end
24. end

During an iterative process, the algorithm removes one state $s$ at a time from the queue (Lines 11) and explores all possible immediate contingencies from that state, identified by $a \in A(s)$ (Lines 12). In particular, for each immediate contingency, SOCCA calculates an expected value of the state from the point of view of the adversary, assuming that all his or her actions in the future are optimal except the immediate next one. The results of these predictions are saved in the buffer list (Lines 13, 14). $R(s,a)$ indicates the benefits of taking malicious action $a$ from state $s$. Once done with all the immediate contingencies from the state $s$, SOCCA sorts the results according to $R(s,a)$ and appends a copy of the ordered buffer to the contingency list output (Lines 16, 17). To update the data structures, SOCCA checks the individual elements in the buffer, and colors the destination states (assuming successful transitions) $s'_{b,s,b,a}$ as gray if they are still white and adds them to the FIFO queue (Lines 18-21). Finally, to prepare for the next iteration, SOCCA clears the temporal buffer list and colors the analyzed state $s$ as black so that it will not be checked again if encountered in future analysis due to the directed loops within the MDP (Lines 24, 25).

It is important to clarify that SOCCA can consider multiple power line outages. In particular, as explained, the proposed framework considers all possible attack paths, each of which often consists of many attack steps, namely several cyber asset vulnerability exploitations as well as one or more maliciously induced physical contingencies such as a power line outage. However, the attack steps cannot grow arbitrarily and are limited by the initial attack point (where the attacker initially resides), i.e., the Internet in our implementations, as well as the network global access policy rules and the system vulnerabilities.

Algorithm 1 can also be employed as an online solution to provide power grid security officers with predictive situational awareness capabilities. Indeed, SOCCA can monitor how future actions by attackers could globally impact the power grid given the current system state. This information is extremely valuable for proactive intrusion prevention systems which reconfigure the system such that the maximum possible damage to the system caused by the attackers’ potential next action is minimized. SOCCA enables the officers to decide which critical components should be monitored more closely in order to detect potential exploitations of known or unknown vulnerabilities. As the algorithm focuses on contingencies originating from malicious cyber attacks, initial contingencies are usually all remote cyber-side vulnerability exploitations. This is because physical devices are almost never directly accessible from a remote machine, unless the attacker has already penetrated deep into the control network. SOCCA takes into consideration possible physical contingencies once states with required set of privileges (compromised host systems) for physical consequences have been reached by the algorithm.

The algorithm considers the path traversal difficulty and the final impact in calculation of the ultimate security index. Therefore, the less likely, yet large impact contingencies may not be fully evaluated during the analysis of Algorithm 1 as they would be ordered further behind while considering the deadline limit of on-line application. To address this point, the algorithm can be updated to always consider the contingencies with high impact and low likelihood. In particular, if the difficulty level of all the cyber-side attacks (i.e., vulnerability exploitations) is set to 0, SOCCA sorts the contingencies only based on their impact level. Consequently, the modified algorithm could partially devote its time to analyze the contingencies based on only their final physical impact.

The description of the contingency screening algorithm so far assumes that its input set including the current system MDP state (set of compromised hosts) and power system state are correct. However, an attacker with the required privilege levels could potentially mislead SOCCA by corrupting the input system state that results in incorrect overall contingency rankings. As a cyber-physical solution against the false data injection attacks against power system state, we have proposed security-oriented cyber-physical state estimation (SCPSE) in [17]. SCPSE takes into account both cyber-side intrusion
VI. Evaluation

In this section, we discuss the implementation of SOCCA and present experimental evaluation results. All our experiments were performed on a 32-bit system with an Intel Core 2 2.16 GHz CPU, 3.00 GB of memory, and the Windows 7 Professional operating system.

Implementations. A unified XML [25] format was used to describe the power system control network topology and network access policy rules (e.g., firewall rules). During the offline phase, SOCCA uses the NetAPT tool [26] to perform a comprehensive security analysis of the access policy rules and to produce the network connectivity matrix according to the control network topology input. The matrix is later translated to the corresponding MDP model. On the power side, we used PowerWorld Simulator [27] to simulate the underlying power system and solve the power flow equations to calculate the cyber-physical security index. In particular, we used the SimAuto toolbox to set up a real-time connection to PowerWorld.

In our experiments, we evaluated SOCCA on a simulated power grid infrastructure. The underlying power system was the IEEE 24-bus reliability test system [28] (Figure 2(a)). The power system consisted of 38 transmission lines, and was monitored and controlled by two control networks with identical network topologies and access control policies. The control network models were built based on topology of a real power control network which is kept anonymous due to non-disclosure agreement. Figure 2(b) shows the topology of a single control network that has 59 nodes, e.g., host systems and firewalls. The first control network monitors and controls buses 1 – 12 in the power system (Figure 2(a)), and the second network monitors and controls buses 13 – 24. In particular, each power bus is monitored and controlled by a single host system in the corresponding control network.

MDP generation. Given the power network topology and the access policy rules, i.e., about 100 firewall rules, SOCCA constructed the network connectivity matrix and generated the corresponding MDP model. It is noteworthy that because the MDP models may not be scalable specially for large-scale power-grid infrastructures, SOCCA makes use of the envelope algorithm [29], where the MDP is generated partially, and hence, not every individual state needs to be enumerated and analyzed. More technically, given the current system state, only reachable states up to some finite horizon are explored and used for the contingency analysis. Figure 3 illustrates a simplified version of the generated MDP in which states with contingencies that are exclusively cyber are drawn in white, while states with physical consequences are in gray. The first number on each state represents its ID. Table I maps each state ID to the IDs of the compromised assets (shown in Figure 2) in each state. As shown on the generated MDP, the attacker initially resides remotely in the internet with no privilege on the power network (MDP’s state 0) and can then traverse different attack paths to access a particular host in the power network. Each MDP edge represents an access (i.e., possibly a vulnerability exploitation) allowed from a source to a destination host in the power network.

Performance. To validate SOCCA’s efficiency on various networks with different sizes and topologies, we measured how long it takes to generate the MDP model for randomly-generated power networks. One important parameter that affects the model generation complexity is the vulnerability factor. This factor is defined as the number of host computers that could be accessed and compromised from a particular host in the power control network. Figures 4(a) and 4(b) show the MDP generation time requirement and the model’s size for vulnerability factor 1, i.e., once the attacker compromises any host system, he or she can always find one other vulnerable machine to compromise. The results were averaged over 1000 runs. As illustrated, for large-scale power networks with 330K host computers, SOCCA analyzed the inputs and generated the MDP model within 24 milliseconds.

For cases with the vulnerability factor of 2, SOCCA generated the corresponding MDP graph within 400 milliseconds for a network with 37 nodes (Figures 4(c) and 4(d)). Figure 5 shows the time taken to solve the security index of individual states. A vulnerability factor of two means that from every host, two other host computers are both accessible (i.e., access control policies allow that) and vulnerable to exploitations. We believe this is very pessimistic and not a very common scenario in industrial control infrastructures due to the extremely strict global access control policies and system security patching within real-world power control networks. The main intention in doing this experiment in the paper was to evaluate, under such pessimistic assumptions, how the increasing network size affects size of the power grid’s MDP model. We continued the experiment for cases with even higher (and much less realistic) vulnerability factors. The model generation for a network with 18 and vulnerability factor of 3 (4) nodes took 74 (5521) milliseconds on average. As expected, for a fixed network topology, the generated model size, and hence the overall performance overhead increases exponentially with the increasing vulnerability factor, which acts as a graph-theoretic branching factor in the MDP model generation procedure (Section III-C). To increase the SOCCA’s scalability for such cases, we implemented the envelope approximation algorithm [29] that, briefly, is based on the control theoretic finite look-ahead optimization technique and considers the next finite set of contingencies while generating the corresponding MDP model. Figures 4(e) and 4(f) show, respectively, how long SOCCA takes, using the
envelope algorithm, to generate the MDP model of the power grid network of different sizes with the vulnerability factor of 4, and the generated MDP model size.

**Metrics.** In our experiments, we pessimistically assumed that all the hosts include security vulnerabilities in order to perform a worst case performance analysis. SOCCA calculates the performance and security indices for individual states in the generated MDP (i.e., shown as second and third number on each state in Figure 3, respectively). In an MDP, there are usually many states with an identical set of physical contingencies that result in equal performance index values. To accelerate the metric calculation and minimize the number of connections to and calculations by PowerWorld, which is a time-consuming step due to the AC power flow solution procedures, SOCCA employs a caching solution to calculate the performance index value for each physical contingency set only once.

**Contingency Ranking.** SOCCA implements Algorithm 1 to rank various security incidents that could occur according to the system’s current state and the generated MDP model once the performance and security indices are calculated for the power grid’s corresponding MDP model. Table II shows the ranked list of cyber-physical contingencies for each state in our case study power grid. It is important to mention that the reported results are for the case in which the attacker has not yet caused any contingency in the power grid, i.e., the current state is \( s_0 = \emptyset \) with ID 0. As shown, the edge \( s_0 \rightarrow s_{14} \) is ranked as the most critical contingency as it allows the attacker to get to the most impactful physical consequence with the least amount of cyber exploitation effort.

**VII. Related Work**

We review the related literature and highlight particular aspects wherein they fall short. Furthermore, we discuss how SOCCA addresses those issues.

There have been several research efforts for computer network contingency analysis, which could be used to enumerate the set of next possible adversarial actions. Static adversary-driven security assessment techniques [30]–[32] explore potential malicious technical actions (contingencies) for every system state before the system goes operational. For instance, ADVISE [30] creates an executable state-based security model of a system and an adversary that represents how the adversary is likely to attack the system and the results of such an attack. As modeling and accurate prediction of the attacker’s behavior are very hard if not impossible in practice, defense-centric security assessment approaches have also been explored recently [33]–[36]. These techniques use manually filled knowledge bases of alert applicability, system configuration, or target importance to associate a context with each alert and to provide security assessment accordingly. The main barriers for real-world deployment of those techniques are the high human involvement required and the lack of awareness for future adversarial actions and social factors. However, static exploration of the whole state space is infeasible in practice due to the state explosion problem and furthermore, physical system contingencies are not accounted for. By design, SOCCA considers attacks consisting of power and/or cyber contingencies and explores a remarkably reduced search space by dynamic generation of the next possible states given the current probabilistic state estimate.

Contingency analysis in power systems has been explored by many researchers in the past (see [37] for a comprehensive survey). The initial efforts were based on first-order performance index sensitivities to rank contingencies, Ejebe et al. [38]. There have been several follow-up attempts to improve the ranking quality by considering higher order sensitivities [39], [40]. Furthermore, there has been an increasing interest
in the analysis of multiple contingencies [41], [42] after the introduction of new NERC standards [43]. Davis et al. [44] propose a linear sensitivity-based approximate measure of how close the power system is brought to islanding by a particular outage contingency. The authors use the metric to categorize various line outages and show that it outperforms similar metrics because of taking care of precisely islanding singularities.

Almost all of the past contingency analysis techniques consider natural incidents to be root causes of the power system contingencies. As a result, they ignore cyber side events and, in particular, contingencies due to deficient or compromised cyber components. To deal with those issues, SOCCA introduces a cyber-physical security formalism that takes into account failure scenarios due to compromised cyber and/or power components.

During the last decade, several researchers have approached the problem of hybrid cyber-physical security modeling for the power-grid from different angles. Mo et al. [4] review a series of security challenges and possible intrusions in power-grid infrastructure. However, a unified modeling framework is not provided. Pasqualetti et al. [6] model a power system under cyber-physical attack as a linear time-invariant descriptor system with unknown inputs, and design a dynamic detection and identification scheme using geometric control theoretic tools. It is not clear from the paper how the cyber network topology affects the way attacks occur in the modeled cyber-physical system. Sridhar et al. [7] review how traditional intrusion tolerance techniques could be applied in cyber-physical settings, and introduce a layered approach to evaluate risk based on the current state of the power-grid. The authors do not discuss how the state is determined, and additionally, accidental failures are not accounted for. Zonouz et al. [8] proposed a framework that fuses uncertain information from different types of distributed sensors, such as power system meters and cyber-side intrusion detectors, to detect the malicious activities within the cyber-physical system. Specifically, they presented a security-oriented cyber-physical state estimation (SCPSE) system, which, at each time instant, identifies the compromised set of hosts in the cyber network and the maliciously modified set of measurements obtained from power system sensors. However, to our knowledge SOCCA is the first framework to consider cyber adversary induced physical contingencies.

VIII. CONCLUSIONS

In this paper, we presented SOCCA, a security-oriented cyber-physical contingency analysis framework that identifies
TABLE II: Ranked (R) List of Cyber-Physical Contingencies (MDP edges)

<table>
<thead>
<tr>
<th>R</th>
<th>Edges</th>
<th>R</th>
<th>Edges</th>
<th>R</th>
<th>Edges</th>
<th>R</th>
<th>Edges</th>
<th>R</th>
<th>Edges</th>
<th>R</th>
<th>Edges</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.74</td>
<td>1</td>
<td>0.1</td>
<td>2</td>
<td>0.196</td>
<td>3</td>
<td>0.743</td>
<td>4</td>
<td>0.746</td>
<td>5</td>
<td>0.747</td>
</tr>
<tr>
<td>10</td>
<td>1.77</td>
<td>11</td>
<td>1.46-142</td>
<td>12</td>
<td>1.46-77</td>
<td>13</td>
<td>1.23-128</td>
<td>14</td>
<td>1.19-122</td>
<td>15</td>
<td>1.19-102</td>
</tr>
<tr>
<td>20</td>
<td>1.48-185</td>
<td>16</td>
<td>1.48-116</td>
<td>17</td>
<td>1.19-141</td>
<td>18</td>
<td>1.19-141</td>
<td>19</td>
<td>1.19-141</td>
<td>20</td>
<td>1.19-141</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>131-116</td>
<td>91</td>
<td>111-38</td>
<td>92</td>
<td>111-97</td>
<td>93</td>
<td>111-122</td>
<td>94</td>
<td>111-122</td>
<td>95</td>
<td>111-122</td>
</tr>
<tr>
<td>108</td>
<td>24-34</td>
<td>108</td>
<td>28-17</td>
<td>109</td>
<td>32-17</td>
<td>110</td>
<td>36-17</td>
<td>111</td>
<td>36-17</td>
<td>112</td>
<td>36-17</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td></td>
</tr>
<tr>
<td>80</td>
<td>1.19-102</td>
<td>81</td>
<td>1.19-102</td>
<td>82</td>
<td>1.19-102</td>
<td>83</td>
<td>1.19-102</td>
<td>84</td>
<td>1.19-102</td>
<td>85</td>
<td>1.19-102</td>
</tr>
</tbody>
</table>

contingencies possible through cyber attacks (e.g., malicious vulnerability exploitations), given the current cyber security state of the power system control network. SOCCA provides the power grid security officers with predictive situational awareness capabilities to assess the global impact of different adversarial actions on the power grid. Thus, it enables operators to decide upon appropriate deployment of proactive intrusion prevention solutions. Our experimental results show that SOCCA complements the traditional power contingency analysis methods that consider physical power component failures due only to accidental failures and other natural causes.

ACKNOWLEDGMENTS

This material is based upon work supported by the Department of Energy under Award Number DE-OE0000997.

REFERENCES


Saman Zonouz is an Assistant Professor in the Electrical and Computer Engineering Department at the University of Miami. He received his Ph.D. in Computer Science from the University of Illinois at Urbana-Champaign in 2011. He has worked on intrusion response and recovery, information flow-based security metrics for power-grid critical infrastructures, online digital forensics analysis and monitorless recoverable applications. His research interests include: computer security and survivable systems, control/game theory, intrusion response and recovery systems, automated intrusion forensics analysis, and information flow analysis-based security metrics.

Charles M. Davis received the B.S. degree in electrical engineering from Louisiana Tech University in 2002 and the M.S. and Ph.D. degrees from the Electrical and Computer Engineering Department at the University of Illinois at Urbana-Champaign. He is currently working at PowerWorld Corporation. His research interests include linear sensitivities, power system analysis, power system visualization, and power system operational reliability.

Katherine R. Davis received the B.S. degree in electrical engineering from the University of Texas at Austin in 2007 and the M.S. and Ph.D. degrees from the University of Illinois Urbana-Champaign in 2009 and 2011. She is a software engineer and senior consultant at PowerWorld Corporation and an Adjunct Assistant Professor in Electrical and Computer Engineering Department at the University of Illinois Urbana-Champaign. Her research interests include data-enhanced power system modeling and analysis and making the grid more robust with respect to bad data.

Robin Berthier is a research scientist at the University of Illinois at Urbana-Champaign, working with Prof. William H. Sanders. Robin graduated from the Reliability Engineering Department at the University of Maryland in 2009. His doctoral dissertation with Prof. Michel Cukier, and focused on the issue of honeypot sensors deployed on large networks. He introduced a new architecture to increase the scalability of high-interaction honeypots, and combined network datasets of different granularities to offer unique attack forensics capabilities. His current research interests include advanced intrusion detection systems and the security of critical infrastructures.

Rakesh B. Bobba is a Research Assistant Professor in the College of Engineering at the University of Illinois, Urbana-Champaign with appointments in Information Trust Institute and Electrical and Computer Engineering Department. His research interests are in the security of distributed and networked systems with a current focus on cyber-physical systems including critical infrastructures such as the power grid and cloud computing. He received M.S. and Ph.D. degrees in Electrical and Computer Engineering from the University of Maryland at College Park in 2007 and 2009, respectively.

William H. Sanders is a Donald Biggar Willett Professor of Engineering, the Interim Head of the Department of Electrical and Computer Engineering, and the Director of the Coordinated Science Laboratory (www.csl.illinois.edu) at the University of Illinois at Urbana-Champaign. He is a professor in the Department of Electrical and Computer Engineering and Affiliate Professor in the Department of Computer Science. He is a Fellow of the IEEE and the ACM, a past Chair of the IEEE Technical Committee on Fault-Tolerant Computing, and past Vice-Chair of the IFIP Working Group 10.4 on Dependable Computing. He was the founding Director of the Information Trust Institute (www.iti.illinois.edu) at Illinois.